MA 551 – Homework 1

Due Feb 12, 2026

Homework Assignment Policy and Guidelines

(a) Homework assignment should be submitted electronically through `Canvas`, and your submission should be combined into ONE PDF file.

(b) Homework assignment should be well organized. It is required that you show your work in order to receive credit.

(c) It's recommended to use R Markdown to write up your homework solutions.

(d) For simulation studies, please refrain from reporting intermediate results; report outputs that are most relevant to the questions.

(e) You may discuss most homework problems with others including your peers and instructor, but you must write up your homework solutions by yourself in order to receive credit. Similarly, you must write your own computer code and obtain your computer output independently.

1. Conduct a simulation study to evaluate the robustness and resistance of a one-sample T test.

(a) Draw a random sample of size $n = 10$ from $N(0, 1)$. Repeat this sampling $S = 1000$ times and for each sample, perform a T test for $H_0 : \mu = 0$ vs. $H_a : \mu \neq 0$ at the $\alpha = 0.05$ level and construct a 95% confidence interval (CI) for μ. Provide the Type I error rate of the T test and the coverage probability of the CI from the 1000 samples.

(b) Repeat (a) but this time let $n = 50$.

(c) Repeat (a) but this time let $n = 200$.

(d) What observations can you make about the results in (a)–(c)?

(e) Repeat (a)–(d), but this time, replace $N(0, 1)$ with χ^2_2 first and then with χ^2_{10}, where χ^2_v refers to a chi-square distribution with v degrees of freedom. The null hypotheses are $H_0 : \mu = 2$ and $H_0 : \mu = 10$ for χ^2_2 and χ^2_{10}, respectively.

2. Implement the bivariate Spearman rank correlation test for independence as a permutation test. The Spearman rank correlation test statistic can be obtained from function `cor` with method = "spearman". Compare the p-value obtained from the permutation test with the p-value reported by `cor.test` on the same samples. Two data examples are considered. In the first example, the samples are drawn from bivariate normal distribution:

```
mu = c(0, 0)
Sigma = matrix(c(1, 0.5, 0.5, 1), 2, 2)
data1 = mvrnorm(10, mu, Sigma)
```

In the second example, the samples are lognormal:

```
data2 = exp(mvrnorm(10, mu, Sigma))
```

3. The `scor` data set in the `bootstrap` package contains test score data on 88 students who took examinations in five subjects. The first two tests (mechanics, vectors) were closed book and the last three tests (algebra, analysis, statistics) were open book. Each row of the data frame is a set of scores $(x_{i1}, \ldots, x_{i5})$ for the i^{th} student.

(a) Use a panel display to display the scatter plots for each pair of test scores. Compare the plot with the sample correlation matrix.

(b) Obtain bootstrap estimates of the biases and standard errors for each of the following estimates: $\widehat{\rho}_{12} = \widehat{\rho}(\text{mec, vec})$, $\widehat{\rho}_{34} = \widehat{\rho}(\text{alg, ana})$, $\widehat{\rho}_{35} = \widehat{\rho}(\text{alg, sta})$, $\widehat{\rho}_{45} = \widehat{\rho}(\text{ana, sta})$.

(c) The five-dimensional scores data have a 5×5 covariance matrix Σ, with positive eigenvalues $\lambda_1 > \ldots > \lambda_5$. In principal components analysis,

$$\theta = \frac{\lambda_1}{\lambda_1 + \ldots + \lambda_5} \tag{1}$$

measures the proportion of variance explained by the first principal component. Let $\hat{\lambda}_1 > \ldots > \hat{\lambda}_5$ be the eigenvalues of $\hat{\Sigma}$, where $\hat{\Sigma}$ is the MLE of Σ.[1]

(i) Compute the sample estimate $\hat{\theta} = \frac{\hat{\lambda}_1}{\hat{\lambda}_1 + \ldots + \hat{\lambda}_5}$ of θ.

(ii) Use bootstrap to estimate the bias and standard error of $\hat{\theta}$.

(iii) Compute 95% percentile and BCa confidence intervals for θ.

Note: the eigenvalues can be obtained from function `eigen`.

[1] Let X be a $p \times 1$ random vector. Its covariance matrix is defined as $\Sigma = E\left[(X - EX)(X - EX)^\top\right]$. Given a sample of n independent observations $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of X where $\mathbf{x}_i$ is a $p \times 1$ vector, the MLE of Σ is $\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^n (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^\top$, where $\bar{\mathbf{x}} = n^{-1}\sum_{i=1}^n \mathbf{x}_i$.